

E L 4 6 5 8 5 4 1 6 8

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION FOR LETTERS PATENT

* * * * *

**MULTI-QUERY DATA VISUALIZATION PROCESSES,
DATA VISUALIZATION APPARATUS, COMPUTER-
READABLE MEDIA AND COMPUTER DATA
SIGNALS EMBODIED IN A TRANSMISSION
MEDIUM**

* * * * *

INVENTORS

NANCY E. MILLER
ELIZABETH G. HETZLER
SUSAN L. HAVRE
KENNETH A. PERRINE
ELIZABETH R. JURRUS
LUCY T. NOWELL

ATTORNEY'S DOCKET NO. BA4-067

1 **MULTI-QUERY DATA VISUALIZATION PROCESSES, DATA**
2 **VISUALIZATION APPARATUS, COMPUTER-READABLE MEDIA AND**
3 **COMPUTER DATA SIGNALS EMBODIED IN A TRANSMISSION**
4 **MEDIUM**

5 **TECHNICAL FIELD**

6 The present invention relates to multi-query data visualization
7 processes, data visualization apparatus, computer-readable media and
8 computer data signals embodied in a transmission medium.

9 **BACKGROUND OF THE INVENTION**

10 This application is related to U.S. Patent No. 6,070,133, entitled
11 “Information Retrieval System Utilizing Wavelet Transform”, issued to
12 M.E. Brewster and N.E. Miller on May 30, 2000 and filed on
13 July 21, 1997, which patent is hereby incorporated herein by reference
14 for its teachings.

15 Some conventional information visualization and retrieval systems
16 provide visualizations related to documents or their attributes by
17 representing documents or a group of documents with graphical symbols.
18 Search techniques for identifying a group of documents or portions of
19 documents relative to some set of search criteria have been developed.
20 Most of these techniques also provide some indicia of relevance for each
21 element harvested by the search.

22 Examples of search techniques and relevancy evaluation tools are
23 discussed, for example, in “Evaluation of a Tool for Visualization of

DRAFT - DO NOT CITE

1 Information Retrieval Results" by A. Veerasamy and N. Belkin, ACM
2 catalogue no. 0-89791-792-8/96/08. This paper discusses a variety of
3 information retrieval strategies and relationships between the search
4 technique and the relevance or interpretation of search results. In
5 general, searches tend to include an initial phase, during which search
6 strategy is "fine-tuned", and a second phase, in which specific items are
7 harvested using the fine-tuned search strategy.

8 In the first phase, interpretation of search results is critical to
9 successful and efficient modification of search strategy in order to try to
10 optimize retrieval of data of particular relevance to a topic of interest.
11 As the amount of data being searched increases, it is increasingly
12 difficult and time-consuming to examine individual documents or portions
13 of documents in order to assess relative relevance to an inquiry. It may
14 also be increasingly difficult to understand relationships between the
15 query, the search tool being employed and the information produced by
16 the search tool. As a result, search results have been organized in a
17 variety of different ways to try to make selected indicia available to the
18 searcher in order to facilitate comprehension of the search results.

19 For example, various types of frequency data may be coupled to
20 specific query elements or search results. As is discussed in the above-
21 noted article, many search engines will display a list of surrogates (e.g.,
22 title, source, author) of the top n-many retrieved items, together with
23 some ranking for each. Such systems do not necessarily provide a clear

1 understanding of why the particular list of items was retrieved, how
2 elements within the list were ranked or how to improve query
3 formulation to arrive at a possibly better set of retrieved data.

4 As the information-handling capacity of data manipulation systems
5 increases, more and more data, running from abstracts to full-text
6 displays, can be provided to the user as the user attempts to focus the
7 search results on the topic of interest. However, this can result in
8 increased search time at the first phase of a search, without necessarily
9 improving the search results or understanding of the relationship between
10 the search criteria and the search results.

11 The types of search tools generally in use allow a relatively
12 complex query to be formulated and are able to provide indicia regarding
13 relevance of search results to components of the query. However, these
14 tools do not lend themselves to simultaneous multiple complex queries
15 and collective interpretation of results from such queries.

16 Accordingly, there is need for visualization systems which provide
17 clear and concise representations of search results that facilitate intuitive
18 understanding of relationships between the search results, the search tool
19 being employed and the queries giving rise to the search results.

20

21

22

23

1 SUMMARY OF THE INVENTION

2 According to one aspect of the present invention, a multi-query
3 data visualization process includes inputting a plurality of query objects
4 into a data processing device and identifying features within each of the
5 plurality of query objects that allow comparison to a body of data stored
6 in a database. The process also includes determining relative
7 relationships between each of the plurality of query objects and the body
8 of data and displaying points along a plurality of rays. Positions of the
9 displayed points correspond to the relative relationships.

10 A second aspect of the present invention provides data visualization
11 apparatus including an image device configured to provide a visual image
12 and digital processing circuitry coupled with the image device. The
13 processing circuitry is configured to input a plurality of query objects
14 and to identify features within each of the plurality of query objects that
15 allow comparison to a body of data stored in a database. The
16 processing circuitry is further configured to determine relative
17 relationships between each of the plurality of query objects and the body
18 of data and to control the image device to depict points corresponding
19 to data from the database along each of a plurality of rays. Positions
20 of the displayed points correspond to the relative relationships.

21 Another aspect of the invention provides computer usable code.
22 The computer usable code is configured to cause digital processing
23 circuitry to identify features of each of a plurality of query objects that

allow comparison to a body of data stored in a database and to determine relative relationships between each of the plurality of query objects and the body of data. The computer usable code is also configured to control an image device to depict points corresponding to data from the database along each of a plurality of rays. Positions of the displayed points correspond to the relative relationships.

A further aspect of the present invention includes a computer data signal embodied in a transmission medium. The signal includes computer usable code configured to input a plurality of query objects into a data processing device and to determine relative relationships between each of the plurality of query objects and a body of data stored in a database. The signal also includes computer usable code configured to control an image device to depict points corresponding to data from the database along each of a plurality of rays. Positions of the displayed points correspond to the relative relationships.

1 **BRIEF DESCRIPTION OF THE DRAWINGS**

2 Preferred embodiments of the invention are described below with
3 reference to the following accompanying drawings.

4 Fig. 1 is a perspective view of an exemplary data visualization
5 apparatus comprising a digital computer, in accordance with an
6 embodiment of the present invention.

7 Fig. 2 is a functional block diagram of exemplary components of
8 the data visualization apparatus of Fig. 1, in accordance with an
9 embodiment of the present invention.

10 Fig. 3 shows an exemplary visual representation corresponding to
11 exemplary data shown upon an imaging medium of an appropriate image
12 device, in accordance with an embodiment of the present invention.

13 Fig. 4 is a graphical representation of an exemplary search results
14 display depicted using the digital computer following reorganization of
15 the data in response to user input, in accordance with an embodiment
16 of the present invention.

17 Fig. 5 shows another exemplary visual representation of the
18 exemplary search results shown in the visual representation of Figs. 3
19 and 4, in accordance with an embodiment of the present invention.

20 Fig. 6 shows an exemplary visual representation corresponding to
21 another form of multi-query based on different forms of similarity to a
22 given graphical object, representing a query or hypothesis, in accordance
23 with an embodiment of the present invention.

1 Fig. 7 is a flow chart illustrating an exemplary process to depict
2 data, in accordance with an embodiment of the present invention.
3

4 **DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS**

5 This disclosure of the invention is submitted in furtherance of the
6 constitutional purposes of the U.S. Patent Laws "to promote the progress
7 of science and useful arts" (Article 1, Section 8).

8 Referring to Fig. 1, a data visualization apparatus 10 is illustrated,
9 in accordance with an embodiment of the present invention. The
10 depicted data visualization apparatus 10 is implemented as a digital
11 computer such as an Ultra 10 elite 3D workstation available from Sun
12 Microsystems Inc. in one exemplary embodiment. Software utilized by
13 the apparatus 10 includes mathematical, analytical and graphical software
14 such as Rogue Wave Software Object-Oriented Libraries including
15 Tools.h++ (Version 7), Math.h++ (Version 6), LAPACK.h++
16 (Version 2), and Analytics.h++ (Version 1) and software graphics
17 package OpenGL™ available from Silicon Graphics, Inc. Other
18 alternatives are possible. The depicted data visualization apparatus 10
19 is configured to operate under a multi-user, multi-tasking operating
20 system, such as UNIX™. Other configurations of data visualization
21 apparatus 10 are provided in other embodiments.

22 As shown, data visualization apparatus 10 includes a plurality of
23 image devices 12, a housing 14 and a user interface 16. Image

1 devices 12 are individually configured to visually depict data such as
2 visual representation 18 described in detail below. Exemplary image
3 devices 12 comprise a monitor 15 and a printer 17. Image devices 12
4 comprise other devices configured to depict data in other embodiments.
5 Exemplary devices of user interface 16 include a keyboard 13 and a
6 mouse 19 as shown.

7 Fig. 2 is a functional block diagram of exemplary components of
8 the data visualization apparatus 10 of Fig. 1, in accordance with an
9 embodiment of the present invention. In particular, housing 14 is
10 configured to house a processor 20, a plurality of storage devices 22 and
11 a network interface 24. In the illustrated configuration, storage
12 devices 22 include memory 26 and disk storage device 28. Storage
13 devices 22 comprise computer usable media configured to store computer
14 usable code and data. Exemplary memory 26 includes random access
15 memory (RAM) and read only memory (ROM). Exemplary disk storage
16 devices 28 include floppy disks and hard disks. Other storage devices
17 such as a CD-ROM device are utilized in other configurations.

18 An exemplary network interface 24 comprises a network interface
19 card configured to couple with an external network such as a public
20 switched telephone network, a packet switched network, such as the
21 Internet etc.

22 Data visualization apparatus 10 is configured to access data and
23 visually depict such data organized as the visual representation 18

(Figs. 1 and 3) with respect to a plurality of query objects and/or events using the image devices 12 in the described embodiment. In the depicted configuration, the visual representation 18 portrays multiple documents or information organized along vectors or rays extending outwardly from a common origin or locus. As used herein, the term "ray" is defined to mean a geometric construct having an origin and a direction, and may correspond to a linear or non-linear construct, such as a spiral, or which may be a directed region of space or volume, such as a half-plane or a curved planar surface. The rays represent the possible variance in relative relationship between the plurality of query objects and the body of data. Documents are illustrated as points spaced apart from the common origin or locus by varying distances. The common origin or locus is representative of the limit of the relative relationships.

The processor 20 comprises digital processing circuitry and is coupled with the image devices 12. The processor 20 is configured to access data from the storage devices 22, the network interface 24 and the user interface 16. The processor 20 is configured to generate the visual representation 18 corresponding to documents, references and/or events within the accessed data as described in detail below. The processor 20 further controls the image devices 12 to depict the visual representation 18 corresponding to the accessed data.

Fig. 3 shows an exemplary visual representation 18 corresponding to exemplary data shown upon an imaging medium 30 of an appropriate image device 12, in accordance with an embodiment of the present invention. The imaging medium 30 is suitable to visually depict the visual representation 18 and in exemplary configurations comprises paper for a printer image device 17 (Fig. 1), a display screen of a monitor image device 15 etc. Other types of imaging media 30 may be used in other embodiments.

Fig. 3 also shows six query objects or inquiries 31-36 grouped about a central point or locus 37. Multiple documents or information each represented by points 38 are organized along rays 41-46 arranged about the central point 37. The rays 41-46 extend outwardly from the common origin or locus 37 where a distance separating each document 38 from the common origin or locus 37 representing the query objects 31-36 represents a degree of similarity or lack thereof with respect to the hypotheses or query objects 31-36. While the rays 41-46 are represented as six rays equiangularly spaced about the locus 37, it will be appreciated that more or fewer query objects 31-36 could be employed, and that the rays 41-46 need not be equiangularly spaced about the locus 37.

The depicted data elements 38 may correspond to the occurrence of particular items (e.g., country names, agricultural products, political movements, legal precedents, technical topics or keywords, image

characteristics etc.) within a body of data, for example. Any type of data may be depicted within the visual representation 18. Types of data that may be analyzed include, for example, images corresponding to tissue samples, micrographs of metal samples, fingerprints or other biometric indicia, or word processing or text-containing files corresponding to legal cases, patent and/or technical publication databases, web documents, audio files of human speech or any other type of data that may be organized into a database.

As used herein, the term "query" is defined to mean an information object to be compared to objects in a database. A query could be one or more words, an image, results of a simulation, a color, a web page, a document, a sound file containing an audio conversation etc. The user is interested in the relative relation between the query and the data in the database. The relationship of interest may include similarity, containment, antithesis, shared attribute etc. The query may be the same kind of entity as the data in the database (for example, using a document as a query to be compared to WWW documents), or it may be different (for example, if the query is a color, and the goal is to find images containing that color). In another example, the query is a scenario and the objects 38 are extracted facts that match elements of the scenario.

The queries may be generated by a single individual or may be generated by multiple people working in a team-oriented or collaborative

1 environment. Thus, for example, Figure 3 might represent a method for
2 exploring how six different people's viewpoints relate to the information
3 in the database.

4 Examples of systems intended to assign numerical surrogates
5 facilitating vector representation for attributes of data within a database
6 in order to promote analysis of bodies of data and data extraction or
7 document retrieval from of bodies of data are described in U.S. Patent
8 No. 5,553,226, entitled "System For Displaying Concept Networks" and
9 issued to Kiuchi et al.; U.S. Patent No. 5,950,196, entitled "System And
10 Methods For Retrieving Tabular Data From Textual Sources" and issued
11 to Pyreddy et al.; U.S. Patent No. 5,659,732, entitled "Document
12 Retrieval Over Networks Wherein Ranking And Relative Scores Are
13 Computed At The Client For Multiple Database Documents" and issued
14 to Kirsch; U.S. Patent No. 5,826,261, entitled "System And Method For
15 Querying Multiple, Distributed Databases By Selective Sharing Of Local
16 Relative Significance Information For Terms Related To The Query" and
17 issued to Spencer, which patents are hereby incorporated herein by
18 reference for their teachings.

19 An exemplary system for carrying out similar sorting and
20 identification with respect to multimedia data is described in U.S. Patent
21 No. 5,873,080, entitled "Using Multiple Search Engines To Search
22 Multimedia Data" and issued to Coden et al., which patent is hereby
23 incorporated herein by reference for its teachings. An example of a

1 system for examining groups of documents and for providing two-
2 dimensional displays related thereto is described in U.S. Patent
3 No. 5,625,767, entitled "Method And System For Two-Dimensional
4 Visualization Of An Information Taxonomy And Of Text Documents
5 Based On Topical Content Of The Documents" and issued to Bartell et
6 al., which patent is hereby incorporated herein by reference for its
7 teachings. Other tools that may be usefully employed include vector
8 space models and statistical natural language processing techniques.

9 Another example of a system for facilitating human interaction with
10 large bodies of information is the Spatial Paradigm for Information
11 Retrieval and Exploration program developed at the Pacific Northwest
12 Laboratory in Richland WA and described, for example, in "Visualizing
13 The Non-Visual: Spatial Analysis And Interaction With Information From
14 Text Documents", published in Proceedings of IEEE '95 Information
15 Visualization, pages 51-58, Atlanta GA, October 1995, available through
16 the IEEE Service Center, and hereby incorporated herein by reference
17 for teachings on information processing and display. The SPIRE™
18 browsing system supports two-dimensional displays of data (e.g., the
19 Galaxy display, similar to Fig. 5, infra) that have been processed to
20 provide feature vector data according to thematic content.

21 The depicted visual representation 18 graphically presents the
22 relationship of each data object 38 in a database to each of the query
23 objects 31-36. The relationship of each data object 38 to a specific

P0000000000000000000000000000000

query object is indicated by the placement of a point representing the data object 38 along a single ray such as 41 corresponding to the query object 31. The proximity of a point along the ray to the locus 37 indicates the strength of the relationship between the query object and the data object represented by the point. In the current embodiment, the closer the point 38 is to the locus 37, the more similar the data object 38 is to the ray's query object. In one embodiment, two-dimensional representations of n-dimensional vectors are prepared using Sammon mapping, as is known in the art. Sammon mapping and other cluster-mapping techniques for representation of n-dimensional vectors in a two-dimensional space are discussed, for example, in U.S. Patent No. 5,897,627, entitled "Method Of Determining Statistically Meaningful Rules" and issued to Leivian et al. and U.S. Patent No. 5,891,729, entitled "Method For Substrate Classification" and issued to Behan et al., which patents are hereby incorporated herein by reference for their teachings.

Additional techniques for mapping data are discussed in U.S. Patent No. 6,031,537, entitled "Method And Apparatus For Displaying A Thought Network From A Thought's Perspective" and issued to Hugh; U.S. Patent No. 6,076,088, entitled "Information Extraction System And Method Using Concept Relation Concept (CRC) Triples" and issued to Paik et al.; U.S. Patent No. 6,026,388, entitled "User Interface And Other Enhancements For Natural Language Information Retrieval System

1 And Method" and issued to Liddy et al.; and U.S. Patent No. 5,576,954,
2 entitled "Process For Determination Of Text Relevancy" and issued to
3 Driscoll, which patents are hereby incorporated herein by reference for
4 their teachings.

5 Query objects 31-36 in accordance with the present invention can
6 take many forms. Query objects 31-36 may correspond to situations
7 where the user does not know much about the expected results, but does
8 know what form a relevant response might take. In this case, the
9 interaction of the user with the database is similar to a conventional
10 search, such as a Boolean keyword search.

11 Query objects 31-36 may represent efforts to browse an information
12 space. In this instance, the user is looking for something, but does not
13 know what the result might look like. Query objects 31-36 may also
14 represent attempts to "reality test" an idea or concept. In this case, the
15 user has a mental model of the content some part of the database, but
16 would like to determine whether the data supports or refutes that the
17 mental model has validity.

18 Examples of types of query objects or hypotheses 31-36 that the
19 user might be interested in may include trying to locate legal precedents
20 for a given fact pattern, trying to locate patents or technical publications
21 relating to a type of device, process or model, searching for information
22 in political speeches, government reports and the like, searching for
23 information regarding chronological developments on a given topic,

1 searching for a subset of images including a some specific type of image
2 or data, searching a series of broadcasts for specific speech patterns,
3 jingles or content or any other form of organized search of a body of
4 data.

5 The processor 20 controls the image device 12 to arrange the
6 visual representation 18 relative to a central locus 37. The locus 37
7 may be provided at other locations relative to the visual
8 representation 18 in other arrangements. Further, the locus 37 may be
9 depicted or not shown at all in particular configurations of the visual
10 representation 18.

11 Fig. 4 is a graphical representation of exemplary search results in
12 visual representation 18 depicted using the digital computer following
13 specification of a relevance threshold 52 in response to user input, in
14 accordance with an embodiment of the present invention. The
15 processor 20 (Fig. 2) is configured to display the rays 41-46
16 corresponding to user-input query objects 31-36 and to determine relative
17 relationships between the points 38 distributed along the rays 41-46 and
18 data stored in the database and to then represent a subset of the data
19 having relevance to the query objects as points 38 distributed along the
20 vectors 41-46 within the relevance threshold 52. In one embodiment, the
21 relevance threshold 52 is represented by a circle or other geometric
22 shape formed about the common origin 37.

1 In one embodiment, the user is able to gauge a probable relevance
2 of data represented by a given point, e.g., point 54, found along one of
3 the rays 41-46, e.g., 43, by noting a distance separating the given object,
4 e.g., that represented by the point 54, from the common origin 37. The
5 object corresponding to the point 54 actually has similar relevance to
6 each of the query objects 31-36 as shown by the arcs 55 coupling the
7 representation of the object 54 on the ray 43 to representations of the
8 object 54 on others of the rays 41, 42 and 44-46. In the example of
9 Fig. 4, the user has requested that the system show all points falling
10 within the relevance threshold 52 for all queries. In this instance, only
11 two objects, represented by the points 54 and 56, meet this criteria.
12 Representations of the object 56 on each of the rays 41-46 are
13 interconnected by arcs 57.

14 In one embodiment, the user may select one of the objects
15 corresponding to the points 54 and 56, e.g., point 54. The selection can
16 be made, for example, using a tactile feedback input device such as a
17 mouse or keyboard (e.g., using arrow keys or the tab key, followed by
18 the enter key). In response to user selection of the given point 54, a
19 display of data relating to the object corresponding to the given point 54
20 is provided. The display may include information such as author,
21 frequency tables for occurrence of selected terms in the query, probable
22 status for the object corresponding to the point 54 vis-a-vis the query 33
23 occurring within the object, confidence factor and the like.

1 For example, in one embodiment, the user may be provided with
2 a text display corresponding to a document represented by the given
3 point 54. In one embodiment, a separate image device displays text
4 corresponding to the document represented by the given point 54. In
5 one embodiment, the user may be provided with a text file corresponding
6 to a portion of a document where the portion has been determined to
7 be that portion of the document that includes reference to a specific
8 theme or idea.

9 In one embodiment, the user may request all objects within the
10 specified distance of all but one of the query objects 31-36, or all but
11 two etc., and to then obtain a display of the ensemble of objects after
12 re-calculation of relative relationships between the query objects 31-36
13 and the collection of objects in the database. In one embodiment, the
14 user may select (e.g., click on) one or more of the queries to turn that
15 query off and to then obtain a display of the ensemble of points after
16 re-calculation of relative relationships between the query objects 31-36
17 and the collection of objects in the database.

18 Fig. 5 shows another exemplary visual representation 58 of the
19 exemplary search results shown in the visual representation 18 of Figs. 3
20 and 4, in accordance with an embodiment of the present invention. In
21 Fig. 5, relative distance represents similarity or lack thereof between
22 distinct points of the representation 58. For example, one method of
23 placing the points (e.g., 38, 31-36, 54) is to use Sammon projection or

DOCUMENT EDITION

1 other multidimensional scaling methods, as described in "Multivariate
2 Analysis" by K.V. Mardia, J.T. Kent and J.M. Bibby, Academic Press
3 Ltd., London, U.K., 1979 (ISBN 0-12-471252-5), which is hereby
4 incorporated herein by reference for its teachings. In one embodiment,
5 the similarity between the query objects and the data in the database is
6 weighted more strongly in determining the positions of points 38 than
7 the similarity among data in the database. In one embodiment, the user
8 may control the weighting scheme, to modify the amount of weighting
9 or to limit it to only some of the query objects 31-36 or some of the
10 database objects. The representations 18 and 58 are linked so that
11 elements (e.g., 31-36, 54, 56) selected in one of the
12 representations 18, 58 also are selected in the other of these
13 representations 18 and 58.

14 Fig. 6 shows an exemplary visual representation 60 corresponding
15 to another form of multi-query based on different forms of similarity to
16 a given graphical object 62, representing a query or hypothesis, in
17 accordance with an embodiment of the present invention. Fig. 6 shows
18 examples of a nearest match 64 interconnected by dashed lines 65 and
19 appearing in each of four different regions 66-72, where each
20 region 66-72 corresponds to an attribute such as black/white mix content,
21 curve content, horizontal component content or spatial frequency content.
22 The object 62 could represent a tissue sample, a metallurgical
23 micrograph, biometric image data or any other type of image data.

1 Fig. 7 is a flow chart illustrating an exemplary process P1 to
2 depict data, in accordance with an embodiment of the present invention.

3 Initially, the processor 20 (Fig. 2) executes a set-up procedure.
4 For example, the processor 20 creates a window having a menu bar
5 and/or a drawing area within the imaging medium of an appropriate
6 image device 12.

7 The process P1 then proceeds to a step S1. In the step S1, the
8 user enters a set of query objects 31-36.

9 In a step S2, the query objects 31-36 are converted to
10 n-dimensional feature data. Conversion to vector data may be carried
11 out using any appropriate algorithm, with the type of algorithm needed
12 being determined in part by the nature of the data forming the query
13 objects 31-36.

14 Next, the processor 20 proceeds to a step S3 to access data
15 objects to be visually depicted by the image device 12. Such data
16 objects typically include references, events or images. In one
17 embodiment, the data consist of entire images or documents. In one
18 embodiment, the data are processed to determine boundaries of portions
19 of data elements, such as documents that are relevant to one or more
20 topics, and the data are broken down into subsets, some of which will
21 be more relevant than others to any given query. In the current
22 embodiment, the feature vectors have already been calculated for the
23 data objects in 38 in the database and are merely accessed in this step.

1 In an alternate embodiment, feature vectors for the data objects 38 could
2 be created or modified based on the queries input in the step S1.

3 In a step S4, the n-dimensional feature vectors of the data objects
4 and the query objects are compared to one another. The step S4
5 determines relationships between each of the data objects 38 in the
6 database and the query objects 31-36.

7 In a step S5, the processor 20 projects the relationships calculated
8 in the step S4 to points along the query rays as seen in Fig. 3. The
9 plurality points along each query ray corresponds to the elements 38.
10 The plurality of query rays corresponds to the query objects 31-36.

11 In a step S6, the processor 20 may optionally reduce the n-
12 dimensional feature vectors of the data objects and the query objects to
13 two- or three- dimensional vectors or points in an alternate projection.
14 In one embodiment, the data object and the query object feature vectors
15 are converted to two-dimensional points using a Sammon mapping as
16 seen in Fig. 5.

17 In a step S7, the processor 20 causes the projected points
18 representing the data objects 38 and the query objects 31-36 to be
19 displayed on one of the display devices 12. In one embodiment, displays
20 of the rays depicting relationships between the data objects and the
21 query objects such as that of Fig. 3 are shown. In one embodiment,
22 displays with alternate projections such as that of Fig. 5 are shown.

1 In a step S8, a relevance threshold is determined. In one
2 embodiment, this results in a display such as that of Fig. 4. In one
3 embodiment, the relevance threshold 52 is set by a user. In one
4 embodiment, the relevance threshold 52 is set according to predetermined
5 characteristics. In one embodiment, the relevance threshold is user-
6 adjustable.

7 In a step S9, a user examines the displayed data. The user may
8 select one or more of the formats illustrated in Figs. 3-5, or may flip
9 from one display type to another.

10 In a query task S10, the process P1 determines when the user
11 wishes to examine attributes of a given point 38 in a display in more
12 detail. When the user wishes to examine attributes of the given point
13 in more detail, control passes to a step S11. When the user does not
14 wish to examine attributes of any points 38 in more detail, or when the
15 user has completed this process, control passes to a query task S12.

16 When the user wishes to examine attributes of a given point 38
17 in more detail, the user may select a limited amount of information
18 (e.g., author, keyword frequency, limited text portions or the like) or
19 more comprehensive information (e.g., a full text version of an object or
20 a detailed image of an object) in the step S11. Control then passes
21 back to the step S9.

22 In the query task S12, the process P1 determines when the user
23 wishes to eliminate one or more of the objects 54 or 56. When the

P03 WO 2012 000000

1 user does not wish to eliminate any elements, the process P1 passes
2 control to a query task S13. When the user does wish to alter or
3 eliminate one or more of the objects such as 54, control passes back to
4 the step S6.

5 In the query task S13, the process P1 determines when the user
6 wishes to alter or remove one or more of the query objects 31-36.
7 When the user wishes to alter one or more of the query objects 31-36,
8 the process P1 passes control to a step S14. When the user does not
9 wish to alter or remove one or more of the query objects 31-36, the
10 process P1 passes control to a query task S15.

11 In the step S14, the user alters or removes one or more of the
12 query objects 31-36. The process P1 then passes control back to the
13 step S2.

14 In the query task S15, the process P1 determines when the user
15 wishes to add one or more new queries. When the user does not wish
16 to add any new queries, the process P1 ends. When the user wishes to
17 add one or more new queries, the process P1 passes control back to the
18 step S1.

19 The processor 20 is configured in one embodiment to adjust
20 control of the data visualization apparatus 12 responsive to input from
21 a user via the user interface 16, via the network interface 24, or other
22 modes. For example, a user may request new data, new time or
23 reference resolution, a curve type for the components, a change in the

1 order of the components or may select or deselect objects with reference
2 to specific ones of the query objects 31-36 or all of them etc. The
3 processor 20 is configured to re-execute appropriate portions of the
4 process P1 responsive to such changes or requests from a user.

5 In compliance with the statute, the invention has been described
6 in language more or less specific as to structural and methodical
7 features. It is to be understood, however, that the invention is not
8 limited to the specific features shown and described, since the means
9 herein disclosed comprise preferred forms of putting the invention into
10 effect. The invention is, therefore, claimed in any of its forms or
11 modifications within the proper scope of the appended claims
12 appropriately interpreted in accordance with the doctrine of equivalents.

13
14
15
16
17
18
19
20
21
22
23